# NON-ADDITIVE COMBINING ABILITIES

## II. RECIPROCALS AND BACKCROSSES

NEIL GILBERT
*John Innes Institute, Bayfordbury, Hertford*
and
J. L. JINKS
A.R.C. Unit of Biometrical Genetics, Dept. of Genetics, University of Birmingham

THIS paper applies the method of Gilbert (1963) to *Nicotiana rustica* data previously reported by Jinks (1954, 1956) and Jinks and Jones (1958). The data consist of flowering dates (in days) and plant heights (ins.) of parents, $F_1$, $F_2$ and backcrosses in each of two years in an $8 \times 8$ diallel cross (including reciprocals). For experimental details, see Jinks (1954, 1956).

Can the concepts of Mendelian genetics be reliably injected into practical analyses of polygenic variation?—the authors cannot agree. Here we attempt a purely statistical description of the data: so the Wolf shall abide with the Lamb. Unhappily we cannot agree about which of us is the Lamb.

## 1. ALGEBRA

This section extends the previous analysis (Gilbert, 1963) to cover reciprocal differences. The analysis is now applied to $k(k-1)$ observations $y_{ij}$. Here $y_{ij}$ might come from the cross between the $i$th parent (as male) and the $j$th parent (as female); or it might come from the backcross of $(i \times j)$ to $(i)$. The method is the same in either case. The model is

$$y_{ij} = m + a_i + b_j + c_i d_j + \text{remainder} \tag{1}$$

where $\Sigma a = \Sigma b = \Sigma c = \Sigma d = 0$ and one arbitrary restriction must be placed on $\Sigma c^2$ or $\Sigma d^2$. If we include selfs $y_{ii}$, the analysis becomes that of a complete $k \times l$ table (here $l = k$) discussed in the previous paper. However, we shall omit selfs, partly because the genetic model of the previous paper suggested that model (1) should not be applied to selfs and crosses simultaneously, and partly because we wish to keep parents ($\equiv$ selfs) and crosses statistically independent. The least squares equations are

$$(k-2) \sum_{j \neq i} z_{ij} d_j = (k-2) c_i \Sigma d^2 - k c_i d_i^2 + \Sigma cd^2 + \frac{d_i \Sigma cd}{k-1} \tag{2a}$$

$$(k-2) \sum_{j \neq i} z_{ji} c_j = (k-2) d_i \Sigma c^2 - k c_i^2 d_i + \Sigma c^2 d + \frac{c_i \Sigma cd}{k-1} \tag{2b}$$

where $z_{ij}$ is the "special combining ability" remaining after fitting additive male and female parental constants according to Yates (1947). These equations reduce to the corresponding equation on p. 67 of Gilbert (1963) when $z_{ij} = z_{ji}$ and $c_i = d_i$ (or $c_i = -d_i$ with the sign of $z_{ij}$ reversed). Here we no longer need to distinguish between the "positive" and "negative" models (1) and (2) of the previous paper. This fact, coupled with the small differences actually observed between $c_i$ and $d_i$, removes one possible objection; namely that one diallel cross might contain a section best analysed by the "positive" model, and another section which requires the "negative". All the plant height data opted for the "negative" model for reciprocal sums (and so $c_i \simeq -d_i$), whereas flowering dates all chose the "positive" (*i.e.* $c_i \simeq d_i$). Since the convergence of the iterative solution of equations (2) is even slower than before, it proved convenient to apply the previous analysis to reciprocal sums and then proceed to a new computer programme for solving equations (2). The sum of squares for multiplicative constants is

$$\Sigma c_i z_{ij} d_j = \frac{1}{k-2} \left[ (k-2)\Sigma c^2 \Sigma d^2 - k\Sigma c^2 d^2 + \frac{(\Sigma cd)^2}{k-1} \right]$$

with $2k-3$ d.f. (Our model places 3 constraints on the values of $c_i$ and $d_i$: or alternatively, the sums of equations (2a), of equations (2b), and of $[c_i(2a) - d_i(2b)]$ are all identically zero.) As pointed out before, the analysis is really concerned with the simultaneous fitting of additive and multiplicative constants to $y_{ij}$. In the analyses presented here, the m.s. for additive combining abilities exceed the m.s. for multiplicative c.a. That is merely a consequence of extracting the additive parameters first, and would be reversed if we fitted multiplicative parameters first.

## 2. 1953 $F_1$ PLANT HEIGHT

As an example we analyse completely one set of data. There is no reason to suppose that the other sets of data, if analysed in similar detail, would behave differently. For each block and for the sum of the two blocks, the additive-multiplicative model has been fitted first to reciprocal sums, and then distinguishing between reciprocals. All figures in this paper are quoted on a per-plant basis.

As Gilbert (1963) pointed out, m.s. for "multiplicative c.a." cannot validly be tested against error: but here they can be tested against each other. Since the model is not additive, s.s. for "block differences" cannot be found directly from block differences, but must be obtained by subtracting s.s. for "block sums" from s.s. for each block separately. Similarly for "reciprocal differences". Now, an additive model applied to block sums must necessarily be exactly equivalent to the same model applied to each block separately (and then summed over blocks). Such equivalence is no longer automatic

TABLE 1

*Analysis of variance of 1953 $F_1$ plant heights*

|  | d.f. | m.s. |
|---|---|---|
| **BLOCK SUMS** |  |  |
| *Reciprocal sums* |  |  |
| Additive c.a. . . . | 7 | 11,516 |
| Multiplicative c.a. . . | 7 | 1461 |
| Remainder . . . | 13 | 201 |
| *Reciprocal differences* |  |  |
| Additive c.a. . . . | 7 | 221 |
| Multiplicative c.a. . . | 6 | 47 |
| Remainder . . . | 15 | 96 |
| **BLOCK DIFFERENCES** |  |  |
| *Reciprocal sums* |  |  |
| Mean . . . . | 1 | 2 |
| Additive c.a. . . . | 7 | 41 |
| Multiplicative c.a. . . | 7 | 59 |
| Remainder . . . | 13 | 33 |
| *Reciprocal differences* |  |  |
| Additive c.a. . . . | 7 | 27 |
| Multiplicative c.a. . . | 6 | 26 |
| Remainder . . . | 15 | 39 |

here, and so it is comforting to notice that the multiplicative c.a. do not differ significantly between blocks. Neither do they show reciprocal differences. In section 4, only the backcross generations are analysed for reciprocal differences.

## 3. ANOMALOUS PARENTS

In two cases (1952 $F_2$ plant height and 1952 backcross flowering date) one very large, negative value of $c$ is counterbalanced by a set of positive values. We shall denote the outstanding value by $c_j$. Such cases arise when the interactions involving parent $j$ are more important than the remainder. In the extreme we have

$$c_i = c + \delta_i/c \qquad\qquad (\text{all } i \neq j)$$
$$c_j = -(k-1)c + \delta_j/c$$

where the $\delta$s, which are small compared with $c$, sum to zero and are chosen so as to predict exactly the observed interactions involving parent $j$, while a purely additive scheme is predicted for the remaining crosses. In such circumstances, the value of $c$ is arbitrary. The case of 1952 $F_2$ plant height is so close to this extreme that it proved impossible to obtain convergence to a final set of $c_i$s, but easy enough to find any number of sets of values, all of which give effectively the same

predictions. One of these sets is quoted in table 5. Parents 1, 2 and 4 show, in general, the largest absolute values of $c$; and Jinks (1954, 1956) interprets the departures from additivity evinced by the crosses involving those parents, as evidence for non-allelic interactions between additive genes. The present analysis is rather different. It not only draws attention to these same parents, but flatly asserts that in 1952 $F_2$ the departures from additivity of parent 4 far outweigh those of any other parent, when examined by this additive-multiplicative model. It is clearly advantageous that the analysis can draw attention to such " anomalous " parents: the technical difficulty over convergence is a consequence of using diallel cross data.

## 4. NUMERICAL RESULTS

" Error " is derived from comparisons between replicates. The remainder m.s. usually exceeds " error ", but only by a factor of about

TABLE 2

*Analyses of variance of $F_1$s*

| | d.f. | m.s. for | | | |
|---|---|---|---|---|---|
| | | plant height | | flowering date | |
| | | 1952 | 1953 | 1952 | 1953 |
| Additive c.a. . . . . | 7 | 4510 | 11,516 | 897 | 3408 |
| Multiplicative c.a. . . . | 7 | 581 | 1461 | 222 | 922 |
| Remainder . . . | 13 | 144 | 201 | 41 | 135 |
| Reciprocal differences . . | 28 | 23 | 117 | 106 | 217 |
| Error . . . . . | 56 | 26 | 37 | 21 | 22 |

TABLE 3

*Analyses of variance of $F_2$s*

| | d.f. | m.s. for | | | |
|---|---|---|---|---|---|
| | | plant height | | flowering date | |
| | | 1952 | 1953 | 1952 | 1953 |
| Additive c.a. . . . . | 7 | 5685 | 10,949 | 1528 | 4504 |
| Multiplicative c.a. . . . | 7 | 771 | 1213 | 232 | 839 |
| Remainder . . . | 13 | 141 | 176 | 47 | 77 |
| Reciprocal differences . . | 28 | 104 | 203 | 126 | 213 |
| Error . . . . . | 167 | 35 | 77 | 32 | 72 |

four, so that the additive-multiplicative model describes the data not perfectly, but well enough for any practical purpose. The genetic model discussed by Gilbert (1963) predicts no reciprocal differences in the backcross multiplicative c.a.; this prediction is not fulfilled.

TABLE 4

*Analyses of variance of backcrosses*

| | d.f. | m.s. for | | | |
| --- | --- | --- | --- | --- | --- |
| | | plant height | | flowering date | |
| | | 1952 | 1953 | 1952 | 1953 |
| Additive c.a. . . . . | 7 | 4955 | 9879 | 1298 | 5930 |
| Additive c.a. × reciprocals . . | 7 | 260 | 745 | 979 | 1924 |
| Multiplicative c.a. . . . | 7 | 494 | 1154 | 211 | 759 |
| Multiplicative c.a. × reciprocals . | 6 | 152 | 249 | 137 | 305 |
| Remainder . . . . | 13 | 97 | 192 | 54 | 56 |
| Remainder × reciprocals . . | 15 | 34 | 64 | 121 | 91 |
| Error . . . . . | 109 | 55 | 66 | 40 | 52 |

TABLE 5

*Values of c for plant height (from $y_{ij} = m + b_i + b_j - c_i c_j$)*

| Parent: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1952 F$_1$ . . | 2·65 | −2·27 | 0·81 | −3·15 | −0·15 | 1·45 | 0·32 | 0·34 |
| 1953 F$_1$ . . | 2·21 | −3·74 | 0·66 | −2·67 | 0·94 | 3·10 | −1·08 | 0·57 |
| 1952 F$_2$ . . | 34·674 | 34·621 | 34·665 | −242·51 | 34·637 | 34·635 | 34·637 | 34·641 |
| 1953 F$_2$ . . | 2·38 | −1·90 | 1·25 | −3·27 | 0·49 | 1·62 | −0·86 | 0·28 |
| 1952 Backcross . | 2·72 | −2·03 | 0·43 | −2·21 | 0·33 | 0·63 | 0·09 | 0·05 |
| 1953 Backcross . | 2·50 | −1·41 | 1·01 | −3·95 | −0·01 | 1·76 | 0·02 | 0·08 |

TABLE 6

*Values of c for flowering date (from $y_{ij} = m + b_i + b_j + c_i c_j$)*

| Parent: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1952 F$_1$ . . | −0·96 | 0·63 | −3·91 | 0·35 | 1·33 | 0·06 | 1·12 | 1·38 |
| 1953 F$_1$ . . | −1·66 | 1·55 | −2·74 | 1·62 | 1·64 | −2·77 | 1·52 | 0·83 |
| 1952 F$_2$ . | −0·27 | 0·86 | −3·24 | −0·46 | 1·21 | −0·16 | 0·95 | 1·11 |
| 1953 F$_2$ . | −2·34 | 0·28 | −0·77 | −0·69 | 3·03 | −1·52 | 0·20 | 1·80 |
| 1952 Backcross . | −3·72 | 0·27 | −0·04 | 0·15 | 1·33 | −0·06 | 0·60 | 1·46 |
| 1953 Backcross . | −2·32 | 0·47 | −1·98 | −0·97 | 2·01 | −0·12 | 1·36 | 1·55 |

The case of 1952 F$_2$ is discussed in section 3. Otherwise the values of $c$ are very similar, indeed much more so than the values of $z_{ij}$ from which they are estimated.

## 5. RELATION BETWEEN GENERATIONS

Since the values of $b$ and $c$ describe the various crosses sufficiently well, we enquire how to predict them from earlier generations. Here the data are unsatisfactory for several reasons. The number of parents is small: regression co-efficients estimated from only eight points must

G 2

be treated with some reserve, however small their estimated statistical errors may be.   The values of $c$ for the cases discussed in section 3 are anomalous.   The values of $b$ and $c$ are themselves highly correlated, and since this is certainly not always the case, any observed relation between generations may be peculiar to the present data.   But, unlike other published cases, the experiments reported here include back-crosses; we shall concentrate on this aspect.

The genetic model mentioned in section 4 (and indeed any genetic model which postulates additive gene effects, and excludes non-allelic interactions) predicts:—

(1) Equal multiplicative c.a. for $F_2$ and backcross reciprocal sums.
(2) Equal additive c.a. for $F_2$ and backcross reciprocal sums.
(3) No reciprocal differences for backcross multiplicative c.a.
(4) Reciprocal difference for backcross additive c.a.
$$= a_i - b_i \text{ in Model (1)}, \quad = \tfrac{1}{2}p_i + \text{constant.}$$

For (1) and (3) see section 4.   The observed values of $(a_i - b_i)$—which, incidentally, are identically equal to the reciprocal differences between the combining abilities of the purely additive model—do not fulfil prediction (4).   In fact, Jinks (1956) uses this criterion to test for non-allelic interactions between additive genes.

TABLE 7

*Regressions of $(a_i - b_i)$ on $\tfrac{1}{2}p_i$*

| | | | |
|---|---|---|---|
| 1952 plant height | . | . | 0·289 ($\pm$0·105) |
| 1953 plant height | . | . | 0·609 ($\pm$0·094) |
| 1952 flowering date | . | . | 0·773 ($\pm$0·132) |
| 1953 flowering date | . | . | 0·769 ($\pm$0·058) |

But prediction (2) is much more accurate:

TABLE 8

*Correlations (6 d.f.) between additive c.a. for $F_2$ and backcross reciprocal sums*

| | | | |
|---|---|---|---|
| 1952 plant height | . | . | 0·979 |
| 1953 plant height | . | . | 0·988 |
| 1952 flowering date | . | . | 0·941 |
| 1953 flowering date | . | . | 0·978 |

The $F_2$ and backcross additive c.a. are, in fact, equal within the limits of error.   The data suggest—but do not prove—that the relation between combining abilities in the same generation ($F_2$ and backcross) is clearer than that between combining abilities in different generations.

## 6. DISCUSSION

What, then, has the analysis achieved?   It has shown once again that the additive-multiplicative model can describe rather accurately the crosses in one generation in terms of two parameters per parent, without postulating any particular genetic model.   The statistical

parameters nevertheless have some (unknown) genetical meaning, since their values in successive generations are always more highly correlated than are the observations from which they are estimated. Since the model is purely statistical, prediction from generation to generation can only be empirical; not enough sets of data have yet been analysed for us to decide whether, and how, such prediction can be usefully made. One criterion will evidently be, whether this statistical prediction improves on the prediction obtainable by fitting a genetic model to the data. Genetic models certainly try to probe deeper into the situation, but they cannot help making numerous simplifying assumptions, since the number of parameters must not exceed the number of observations. The choice is between a statistical model with no theoretical basis, and a genetic model with (possibly) too much.

## 7. SUMMARY

An additive-multiplicative statistical model is extended to include reciprocal differences. The model describes the crosses in any one generation remarkably well; but prediction from one generation to the next remains uncertain.

## 8. REFERENCES

GILBERT, N. 1963. Non-additive combining abilities. *Genet. Res., 4,* 65-73.

JINKS, J. L. 1954. The analysis of continuous variation in a diallel cross of *Nicotiana rustica* varieties. *Genetics, 39,* 767-788.

JINKS, J. L. 1956. The $F_2$ and backcross generation from a set of diallel crosses. *Heredity, 10,* 1-30.

JINKS, J. L., AND JONES, R. M. 1958. Estimation of the components of heterosis. *Genetics, 43,* 223-234.

YATES, F. 1947. The analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity, 1,* 287-301.